

# ***Clinical Indicators - Data Quality Filters (Calendar Years 2000-2002)***

## ***Technical specification***

### **Table of Contents**

1	Purpose.....	2
2	Data Quality: Coverage.....	2
	2.1 Assessing Coverage: Primary Care Trust.....	2
	2.2 Assessing Coverage: Acute/Mental Illness Trusts .....	3
3	Data Quality: Completeness .....	4
	3.1 General definition.....	4
	3.2 Filters included.....	4
	3.3 Fields included in the Indicators.....	4
	3.4 Methodology for assessing missing/invalid data .....	6
	3.5 Data quality validity marks .....	7
4	Data Quality: Clinical coding of death episodes .....	7

## 1 Purpose

The Data Quality Filters (DQF) are used to evaluate the quality of the records used to produce the Clinical Indicators (CIs). Each organisation is given a mark to indicate the quality of its data and to determine whether a CI will be included in the publication of the star ratings. There are three elements included in the construction of the CI DQF: coverage, completeness, and death coding.

Coverage tests for a discrepancy between the number of Hospital Episodes Statistics (HES) records received and the independent activity count supplied by the Trust, or with previous years if no activity count was submitted. Data completeness tests for the proportion of complete fields included in the construction of the data filters. Lastly, the death code indicator applies to two indicators (CIs 1625 and 1024) and assesses whether there is a primary diagnosis for all episodes resulting in a death. These three elements will be explained further in the following sections.

One indicator (CI 1024) is based on three calendar years of data, “deaths within 30 days of a heart bypass operation”. A trust must pass the data quality tests for all three years in order to be included.

## 2 Data Quality: Coverage

Coverage relates to the percentage of missing HES records. A measure of coverage is calculated for each organisation for each year. Coverage is assessed based on different criteria, depending on whether an organisation is a Primary Care Trust (PCT) or Acute/Mental Health Trust.

Organisations that fail the pre-defined levels of coverage for a particular year are excluded from all CIs for that year. Please note that CI Indicator 1024 (1C) – “deaths within 30 days of CABG”; is based on three years of data. Therefore trusts must pass the data coverage checks, for all three years to be included.

### 2.1 Assessing Coverage: Primary Care Trust

For each PCT of responsibility, HES coverage is assessed by comparing the level of inpatient activity (the total number of Finished Consultant Episodes (FCE's)<sup>1</sup>) to the GP population (Population assigned to GP area (PCT of responsibility)). The number of units of activity per 100 head of population is calculated. The cut off point is 15 units of activity per 100 head of population. Table 1 outlines the criteria for assessing coverage for PCTs.

**Table 1 Criteria for assigning data quality coverage marks to PCT of responsibility**

Quality mark	Description	Criteria
+	Adequate data quality coverage	15 units of activity or more per 100 head of population
Q	Low data quality – excluded because of low record coverage	Less than 15 units of activity per 100 head of population

<sup>1</sup> The number of episodes where epistat = 3 (finished) and classpat = 1,2 or 5 (ordinary admissions, day cases and mothers and babies).

## 2.2 Assessing Coverage: Acute/Mental Illness Trusts

For each Trust, HES coverage is assessed by comparing the level of inpatient activity (the total number of FCEs) with an independent figure of Trust activity. Table 2 below illustrates the independent figure of Trust activity used for each calendar year.

**Table 2 Independent comparator of Trust activity by year**

Datayear	Comparator		
	1 <sup>st</sup> comparator	2 <sup>nd</sup> comparator if 1 <sup>st</sup> unavailable	3 <sup>rd</sup> comparator if 1 <sup>st</sup> and 2 <sup>nd</sup> unavailable
2000	Activity counts <sup>2</sup> (2000/2001)	KP70 (for 1999/2000)	1999 No. of FCE's on HES
2001	Activity counts (2000/2001)	2000 No. of FCE's on HES	KP70 (for 1999/2000)
2002	Activity counts (2001/2002)	Activity counts (2000/2001)	2001 No. of FCE's on HES

If the ratio of the smallest figure (either the comparator or the HES count) divided by the largest figure (either the comparator or the HES count) is lower than 80%, then the Trust fails the data coverage check for that particular year. Table 3 outlines the method used to assign coverage marks to Trusts. Table 4 displays the list of indicators and whether they apply to a PCT or Trust.

**Table 3 Criteria for assigning data quality coverage marks to Trusts**

Quality mark	Description	Criteria
+	Adequate data quality coverage	At least 80% coverage
Q	Low data quality – excluded because of low record coverage	Less than 80% coverage

**Table 4 Indicators by Trust and PCT of responsibility**

Indicator	Trust	PCT of responsibility
1625 (1a) <sup>3</sup>	√	√
1024 (1c)	√	√
1028/1029 (4a/4e)	√	√
4566 (7)		√
1026/1027 (4f/4g)	√	√
3567/3566 (9e/9f)	√	√

<sup>2</sup> Activity counts: Trusts are sent details of the numbers of FCE's on HES and asked whether these reflect the number on their own database, to check that all data has passed through to the HES system via the NHS clearing service. If the figures do not tally, the Trusts provide a count of the number of FCE's that *should* be on HES.

<sup>3</sup> The codes in brackets are the previous indicator codes.

### 3 Data Quality: Completeness

#### 3.1 General definition

Data quality validity indicates the quality of the HES records.

##### 3.1.1 Four data quality validity groups

There are some indicators that can be grouped together in that they are assessed for data quality in the same manner. Table 5 shows which data quality filters apply to the grouped indicators (i.e., CIs 1027 and 1028).

**Table 5 CI's by each DQF CI number is followed by historical reference number**

Grouped	CI
	1625 (1a) deaths within 30 days of surgery – non-elective admissions
	1024 (1c) deaths within 30 days of heart by pass operation
√	1028 (4f) emergency readmissions – fractured hip 1029 (4g) emergency readmissions – stroke
√	4566 (7) Bronchopneumonia and pneumonia (only applicable for PCTs) 1026 (4a) emergency readmission to hospital following discharge 1027 (4e) emergency readmission of children to hospital following discharge
√	3567 & 3566 (9e/f) readmission to hospital following a psychiatric discharge

#### 3.2 Filters included

Different filters are used in each indicator i.e. different selections of data – see Table 6 below.

**Table 6 Filters by DQF Indicators:**

	1625 (1a)	1024 (1c)	1027, 1028 (4f/4e)	4566 (7)	1026, 1029 (4a/4g)	3567, 3566 (9e/f)	Cove rage	Death coding
Episode type								
General (1)	√	√	√	√	√	√	√	√
Delivery (2)	√	√	√	√	√	√	√	√
Birth (3)							√	
Episode status								
Finished (3)	√	√	√	√	√	√	√	√
Unfinished (1)			√		√	√		
Patient classification								
Ordinary admissions (1)	√	√	√	√	√	√	√	√
Day cases (2)			√		√	√	√	√
Mothers and babies (5)	√	√	√		√	√	√	√
Discharge method								
Deceased								√

#### 3.3 Fields included in the Indicators

Each DQF for the clinical indicators examines a different set of fields. For each DQF indicator, each episode is scanned for missing or invalid data in any of the included fields. If any of the fields in that episode contain missing or invalid data, the whole episode is marked as having invalid/missing data. Table 7 summarises the variables included in the construction of the indicators.

**Table 7 Summary of fields and filters applied to each DQF indicator**

DQF Indicator	Sex	Epistart <sup>4</sup>	Epiorder <sup>5</sup>	Epiend <sup>6</sup>	Admimeth <sup>7</sup>	Admidate <sup>8</sup>	Disdate <sup>9</sup>	Dismeth <sup>10</sup>	Disdest <sup>11</sup>	Startage <sup>12</sup>	Oper_1-12 <sup>13</sup>	Diag_1 <sup>14</sup>	Admisorc <sup>15</sup>	Resstha <sup>16</sup>	Mainspef <sup>17</sup>	Tretspef <sup>18</sup>
1028,1029 (4f/4g)	√	√	√	√	√	√	√	√	√	√		√	√	√	√	
4566 (7)	√	√	√		√					√		√		√		
1625 (1a)	√	√	√	√	√	√	√	√	√	√	√	√	√	√		
1024 (1c)	√	√	√	√	√	√	√	√	√	√	√		√	√		
1026,1027 (4a/4e)	√	√	√	√	√	√	√	√	√	√		√	√	√	√	√
3567,3566 (9e/f)	√	√	√	√	√	√	√	√	√	√			√	√	√	√

<sup>4</sup> Episode Start date

<sup>5</sup> Episode order

<sup>6</sup> Episode end date

<sup>7</sup> Method of admission

<sup>8</sup> Date of admission

<sup>9</sup> Discharge date

<sup>10</sup> Method of discharge

<sup>11</sup> Discharge destination

<sup>12</sup> Age at start of episode

<sup>13</sup> Operative procedure

<sup>14</sup> Primary diagnosis

<sup>15</sup> Source of admission

<sup>16</sup> Strategic health authority of responsibility

<sup>17</sup> Main specialty

<sup>18</sup> Treatment specialty

### 3.4 Methodology for assessing missing/invalid data

The method for assessing whether a particular field contains missing or invalid data is summarised in Table 8 below.

**Table 8 Methodology for assessing missing/invalid data by field**

Field	Missing/invalid data	Criteria
Sex	0,9	
Epistart (Episode start date)	Null	
Epiorder (Episode order)	99,98	
Epiend (Episode end date)	Null or epiend<epistart	Status episode = '3' (Finished)
Admimeth (Method of Admission)	99,98	
Admidate (Admission date)	Null	
Disdate (Discharge date)	Null	End of spell = 'Y' <sup>19</sup>
Dismeth (Method of Discharge)	9	End of spell = 'Y'
Disdest (Discharge destination)	99	End of spell = 'Y'
Startage (Age at start of episode)	Null	
Oper1_12 (Operative procedures #1-12)	&	Status episode = '3' (Finished)
Diag_1 (Primary diagnosis)	R69X, R95-R99, or 6th character = 'A' <sup>20</sup>	Status episode = '3' (Finished)
Resstha mapped (Strategic health authority of residence)	Y	
Admisorc (Source of Admission)	98,99	
Mainspef (Main specialty)	&	
Tretspef (Treatment specialty)	&	

The total number of episodes is calculated for each organisation by each DQF indicator. The total numbers of episodes with missing/invalid data are counted for each DQF indicator. The percentage of episodes with missing/invalid fields is produced for each DQF indicator by organisation, for each year.

<sup>19</sup> This rule applies because discharge information (dismeth, disdate and disdest) is presented on discharge episodes only.

<sup>20</sup> A diagnostic code with an 'A' in the sixth character position identifies an asterisk code, which should not occur in the primary diagnosis.

### 3.5 Data quality validity marks

Trusts that do not have adequate data quality for one indicator may pass on another indicator. Please note that CI 1024 (1C) (“deaths within 30 days of a heart bypass operation”) is based on three years data. Therefore organisations must pass the data quality check for all three years to be included. See Table 9 below for details:

**Table 9 Criteria for assigning data quality validity marks to organisations: DQF indicators**

Quality mark	Description	Criteria
+	Adequate data quality validity	Less than 20% FCE's missing/invalid
Q	Low data quality - excluded because of low record quality	20% or more FCE's missing/invalid

## 4 Data Quality: Clinical coding of death episodes

A further data quality validity component checks the quality of clinical coding of death episodes. This component checks whether or not a primary diagnosis is present for each discharge ending in death. The percentage of records without a diagnosis must be lower than 20% for indicators to be published (CIs 1625 and 1024). Table 10 below indicates the criteria involved in assigning quality marks for the death codes. Only two indicators (CIs 1625 and 1024) examine death coding as part of the data quality component.

**Table 10 Criteria for assigning data quality validity mark to organisations: the deaths diagnostic coding indicator**

Quality mark	Description	Criteria
+	Adequate clinical coding of deaths	Less than 20% primary diagnoses for FCE's missing/invalid
Q	Low data quality - excluded because of poor clinical coding of deaths	20% or more primary diagnoses for FCE's missing/invalid